SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 16, No. 9, pp. 1785–1788, 1997

Printed in the USA
0730-7268/97 $6.00 + .00

EXTRAPOLATION FACTORS FOR SMALL SAMPLES OF PESTICIDE TOXICITY DATA: SPECIAL FOCUS ON LD50 VALUES FOR BIRDS AND MAMMALS

ROBERT LUTTIK* and TOM ALDENBERG
National Institute of Public Health and the Environment, P.O. Box 1, 3720 BA Bilthoven, The Netherlands

(*Received* 22 *April* 1996; *Accepted* 8 *October* 1996)

Abstract—For the registration procedure of pesticides it is necessary to have specific information and testing data in order to conduct an ecological hazard/risk assessment. The hazard/risk assessment for acute exposure is usually based on a quotient method, where the estimated environmental concentration is compared with the lowest available 50% lethal dose (LD50) value. Generally there are only one or two LD50s available for birds and mammals, and an underestimation of the potential hazard/risk could be a real possibility. In this article, we propose to estimate a hazardous dose for 5% of the species (HD5) for LD50 data, corresponding with the hazardous concentration for 5% of the species for the no-observed effects concentration (NOEC) data. In addition we describe a method for calculating safety factors that can be used in the case of small sample sizes, especially those smaller than 4 and including $n = 1$. The safety factors to be applied to the geometric mean of the LD50s for a median estimate of the HD5 of birds and mammals are 5.7 and 3.8, respectively. The safety factors for the 95% confidence limit of the HD5 of birds for $n = 1$, 2, and 3 are 33, 20, and 16, respectively. For mammals these safety factors are 15, 10, and 8.

Keywords—Hazard/risk assessment Extrapolation Safety factors Birds and mammals

INTRODUCTION

The evaluation of the possible environmental hazard for birds and mammals arising from the use of an agricultural pesticide is usually based on the lowest available LD50 (the statistically derived single dose of a substance that can be expected to cause death in 50% of birds or mammals when administered by the oral route). Generally there are only one or two LD50s for birds available (the mallard and quail) and one for mammals (the rat) [1–3]. Therefore, an underestimation of the potential hazard is a real possibility.

A general approach toward extrapolating laboratory (no-observed effects concentration, NOEC) data to an acceptable concentration in the field is to estimate the HC5, the hazardous concentrations for 5% of the species [4–7]. The NOEC is the maximum treatment level in a test that produces no adverse effects. Extrapolations factors k_n have been tabulated for the logistic statistical distribution [7] and the normal distribution [6]. There are extrapolation factors for median estimates of the HC5 and for one-sided 95% left confidence limits of the HC5. Although this procedure formally works from $n = 2$ onward, in practice the method is not used for $n = 2$ and 3. In the latter case, the extrapolation factors are large, especially in the case of the left confidence limit, to account for the uncertainty due to estimating means and standard deviations from such a small sample size. When only one NOEC is available, the method cannot be applied. In this article, we apply the HC5 methodology for NOECs to LD50s of birds and mammals and introduce an HD5 (hazardous dose) analogous to the HC5 for these data. In addition, we explicitly address the problem of very small sample sizes, especially those smaller than 4 and including $n = 1$. This can be done by using information on the standard deviation of the species toxicity distributions for other toxic substances for which data on four or more species are available.

METHODS

For the collection of acute oral toxicity data (LD50) for birds and mammals the literature present at the Centre for Substances and Risk assessment of the National Institute of Public Health and Environment were used. The LD50 values from review articles and handbooks were used as such, because acute oral studies have already been carried out for a long time and are standardized to a great extent. Only when strong indications were present about the unreliability of a study, these data were not accepted. For these cases where it was not clear whether the presented data were separate data or a range (e.g., 10–40 mg/kg body wt.), the data were treated as separate single values.

The values were prepared in the following way for the extrapolation method:

—if more than one LD50 value was available for a species a geometric mean value was calculated;
—if in a set of available LD50s for a certain species a greater or lower than value was present, this value was not used when the value was inside the range of values and was used when the value was outside the range;
—if in a set of available LD50s for a compound for a particular species only a greater or lower than value was present, this value was only used as such when this value was outside the range of all the other values.

Only compounds with data on four or more different species were used for this research (raw data on the selected compounds are reported in Luttik and Aldenberg [8]).

In order to derive information on the spread in sensitivity of different species for well-studied toxicants, the first step is

* To whom correspondence may be addressed.
Presented in the Symposium on Wildlife Ecotoxicology, 1995 SETAC-Europe Meeting, Copenhagen, Denmark, June 25–28.

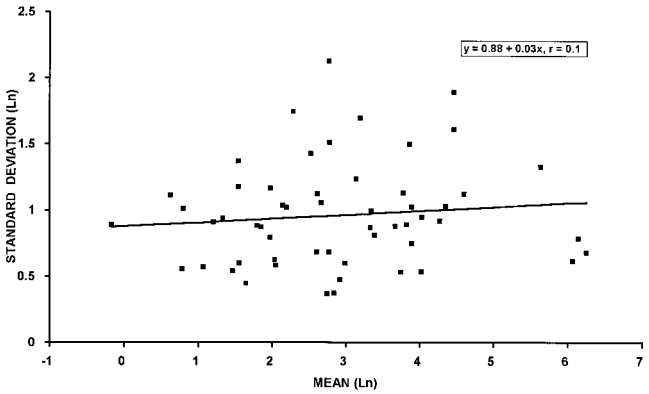

Fig. 1. Averages and standard deviations of ln-transformed LD50s (mg/kg body wt.) for 55 compounds for birds.

to investigate whether the sample standard deviations s_{ni} for these toxicants are independent of the means of the species ln(LD50) distributions. This is illustrated in Figures 1 and 2. The correlation coefficient for birds is 0.106 ($p = 0.441$, two-sided), and for mammals this coefficient is -0.226 ($p = 0.062$, two-sided). Meaning that, at the 5% level the correlation for both groups does not differ significantly from zero. Hence, it seems reasonable to assume that the respective standard deviations are independent of the respective means.

An estimate of a pooled or averaged estimate of the variance [9], when m datasets are available, is

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \ldots + (n_m - 1)S_m^2}{n_1 + n_2 + \ldots + n_m - m} \quad (1)$$

where $S_1^2, S_2^2, \ldots, S_m^2$ are the respective sample variances of the ln(LD50) datasets for m different toxicants. Note that with one toxic substance, i.e., $m = 1$, the one-sample standard deviation estimate is obtained.

From now on it is assumed that s_p is an estimate of the standard deviation of species sensitivities, to be used when nothing is known about the variation in species sensitivity for the toxic substance under study.

Let us consider a sample of n ln(LD50) data, where n may be small, e.g., below 4, including the case $n = 1$, and suppose the data came from a logistic distribution with parameters α and β [7]. Under these conditions, the mean and standard deviation are defined as $\mu = \alpha$ and $\sigma = \beta\pi/\sqrt{3}$. When μ and σ are known, $\ln(\text{HD5}) = \mu - 1.62\sigma$ [7].

Hence, assuming that σ is given through the pooled estimate

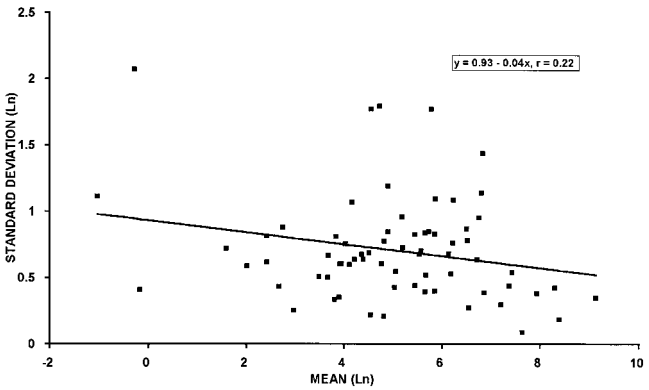

Fig. 2. Averages and standard deviations of ln-transformed LD50s (mg/kg body wt.) for 69 compounds for mammals.

s_p, defined above, as calculated from other data sets, ln(HD5) may be estimated by the statistic

$$\ln(\text{HD5}) = \bar{x} - 1.62\sigma \quad (2)$$

where $\bar{x}$ is the mean of the sample.

This statistic is logistically distributed for $n = 1$ and becomes increasingly normally distributed for $n > 1$. The mean is the true ln(HD5) and the standard deviation is approximately $\sigma/\sqrt{n}$. This procedure overestimates as well as underestimates the true ln(HD5) by 50% and therefore is also a median estimate.

The 5th and 95th percentile of the $\bar{x} - 1.62\sigma$ are located at $\mu - 1.62\sigma \pm 1.64\sigma/\sqrt{n}$, with 1.64 being the Z-value of the standard normal distribution at the 5th percentile. Therefore, the distribution of

$$\bar{x} - (1.62\sigma + 1.64\sigma/\sqrt{n}) \quad (3)$$

underestimates the true ln(HD5) by 95% and is a left confidence limit of ln(HD5). In the original concentration unit the median and left/right confidence limits become

$$\text{HD5} = \exp(\bar{x} - 1.62\sigma) \quad (4)$$

and

$$\text{HD5} = \exp[\bar{x} - (1.62 \pm 1.64/\sqrt{n})\sigma] \quad (5)$$

The safety factors (SF) are defined according to Kooijman [4] as the geometric mean of the original data divided by one of the estimates of the HD5

$$\text{SF} = \exp(\bar{x})/\text{HD5} \quad (6)$$

Thus, one can consider a median estimate SF as

$$\text{SF50} = \exp(1.62\sigma) \quad (7)$$

and a left confidence limit SF as

$$\text{SF95} = \exp[1.62 + (1.64/\sqrt{n})\sigma] \quad (8)$$

RESULTS AND DISCUSSION

For birds 55 compounds (see Appendix) were found with LD50 data on 4 or more species (up to 34 species) and for mammals 69 compounds (see Appendix) with data on 4 or more species (up to 14 species). With these data the standard deviations of the ln-transformed LD50s for each compound can be calculated. The results are presented in Figure 1 for the birds and in Figure 2 for the mammals. The calculated s_p values (with Eqn. 1) are 1.071 and 0.829 for birds and mammals, respectively. With these s_p values SFs can be calculated (Table 1) when one has LD50 values for a small number of different species

In Table 2 the extrapolation constants of this method are tabulated for n together with those of Table 3 in Aldenberg and Slob [7]. The extrapolation constants for the HD are less severe than those for the HC, because in the latter case the standard deviation is estimated from the sample itself, while in the former case the standard deviation is calculated from an external data set that can be considered as a fixed value. For large samples the extrapolation constants are becoming identical (i.e., 1.62).

One of the premises when carrying out the extrapolation method is that the distribution of the input data should be a logistic distribution. This was tested with the goodness-of-fit test of the E_TX program of Aldenberg [10], which is based on

Table 1. Safety factors to be used for the hazard/risk assessment based on LD50s for birds and mammals

Number of LC50s	Safety factors for birds		Safety factors for mammals	
	SF50[a]	SF95[b]	SF50	SF95
1	5.7	32.9	3.8	14.9
2	5.7	19.6	3.8	10.0
3	5.7	15.6	3.8	8.4
4	5.7	13.7	3.8	7.6
5	5.7	12.4	3.8	7.0
6	5.7	11.6	3.8	6.7
7	5.7	11.0	3.8	6.4
8	5.7	10.6	3.8	6.2
9	5.7	10.2	3.8	6.0
10	5.7	9.9	3.8	5.9

[a] Median estimates of safety factors.
[b] One-sided left confidence limit.

the Kolmogorov–Smirnov test statistics according to D'Agostino and Stephens [11].

The results of these goodness-of-fit tests show that at the 5% significance level in 84 and 93% of the cases, for birds and mammals, respectively, the hypothesis that the data are derived from the logistic distribution is not rejected. These results support the assumption that data on sensitivity of birds or mammals for a certain compound come from a logistic distribution, and hence the application of Equations 4 and 5 is allowed.

Although the method is described for calculating safety factors to be used for small samples of LD50s for birds and mammals, the method is generally applicable and can be used for example for small samples of NOECs and LC50 values. In addition SFs can be calculated for different groups of chemicals (see Appendix). For birds the pooled ln standard deviations of the LD50s for carbamates ($n = 9$) and organophosphorous ($n = 28$) compounds are 0.95 and 1.01, respectively. These pooled ln standard deviations are somewhat lower than the pooled ln standard deviation of 1.07 of the total sets of compounds. For mammals the pooled ln standard deviations for carbamates ($n = 13$) and organophosphorous ($n = 23$) compounds are 0.69 and 0.78, respectively. They are also somewhat lower than the pooled ln standard deviation of 0.83 of the total set of compounds. Lower pooled ln standard deviations are also expected for nonpesticides (for instance narcotic compounds). This could not be tested because more than 1 or 2 LD50s are hardly ever available for these types of chemicals. The SF for $n = 1$ based on the HD5 for birds is approximately 1.5 to 2 times as large as the one for mammals (5.7 vs 3.8 and 33 vs 15). The following question can be raised: is this difference between mammals and birds a real difference? First, the difference between these SFs is a reflection of the difference between the pooled ln standard deviations for birds and mammals, 1.071 and 0.829, respectively. The mean ratio between the lowest and highest available toxicity data (see Appendix) for birds within the compounds is 117 (range 4–1,280) and for mammals 34 (range 2–842). Second, the data set for mammals could be biased, because in contrast with the available data for birds, the information on mammals is to a great extent based on representatives of only two orders: rodents and lagomorphs, 56% and 17%, respectively. But the pooled ln standard deviation for compounds with less than 60% representatives of the rodents and lagomorphs is even smaller than the pooled ln standard deviation for the total data set. These indications lead to the conclusion that the difference between the SF, i.e., the difference between the pooled ln standard deviations is a real difference (the two mean standard deviations are significantly different, $p = <0.001$).

Table 2. Extrapolation constants for the calculation of median estimates (50%) and one-sided left confidence limits (95%) for the logarithmic hazardous concentration (HC) and hazardous dose (HD) for 5% of the species on the basis of logistic distribution

Sample size (n)	Extrapolation constants			
	HC5(50)	HD5(50)	HC5(95)	HD5(95)
1	—	1.62	—	3.26
2	2.49	1.62	27.70	2.78
3	2.05	1.62	8.14	2.57
4	1.92	1.62	5.49	2.44
5	1.85	1.62	4.47	2.35
6	1.81	1.62	3.93	2.29
7	1.78	1.62	3.59	2.24
8	1.76	1.62	3.37	2.20
9	1.75	1.62	3.19	2.17
10	1.73	1.62	3.06	2.14
20	1.68	1.62	2.49	1.99
30	1.66	1.62	2.28	1.92
50	1.65	1.62	2.10	1.85
100	1.63	1.62	1.95	1.78
200	1.63	1.62	1.85	1.74
500	1.63	1.62	1.76	1.69
∞	1.62	1.62	1.62	1.62

RECOMMENDATIONS

It is recommended that SFs of the one-sided 95% confidence limit of the HD5 based on the pooled ln standard deviations be used when only 1, 2, and 3 LD50s are available. For $n \geq 4$ the Aldenberg and Slob method [7] can be used. The HD5 values for the compounds that are used for the estimation of the s_p value are presented in the Appendix.

We propose to use the more conservative method (95% confidence limit) for the extrapolation, because the pooled ln standard deviation that is used is not a worst case value. Hence the probability to overpredict ln(HD5) should be small (5%). When there are indications that the estimated standard deviation (according to Aldenberg and Slob [7]) for data sets with $n \geq 4$ could be unreliable, one could consider using SFs based on the pooled ln standard deviation. When there are indications that the available LD50 could be derived from a test with a sensitive species one could consider using the SF for a median estimate of the HD5.

Acknowledgement—The authors thank J. Hoekstra and W. Slob for their statistical advice and for comments on the manuscript.

REFERENCES

1. **Pesticides Bureau.** 1991. Guidelines for the submission of applicants for registration of pesticides. Part H. Toxicity to organisms in the environment. Wageningen, The Netherlands.
2. **U.S. Environmental Protection Agency.** 1994. Pesticides reregistration rejection rate analysis. Ecological effects. EPA 738-R-94-035. Washington, DC.
3. **European Economic Community.** 1996. Concerning the placing of plant protection products on the market. L 65/20. Official Journal of the European Communities.
4. **Kooijman, S.A.L.M.** 1987. A safety factor for LC50 values allowing for differences in sensitivity among species. *Water Res.* **21**:269–276.
5. **Straalen, N.M. van** and **C.A.J. Denneman.** 1989. Ecotoxico-

logical evaluation of soil quality criteria. *Ecotoxicol. Environ. Saf.* **18**:241–251.
6. **Wagner, C.** and **H. Løkke.** 1991. Estimation of ecotoxicological protection levels from NOEC toxicity data. *Water Res.* **25**:1237–1242.
7. **Aldenberg, T.** and **W. Slob.** 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicol. Environ. Saf.* **25**:48–63.
8. **Luttik, R.** and **T. Aldenberg.** 1995. Extrapolation factors to be used in case of small samples of toxicity data (with a special focus on LD50 values for birds and mammals). RIVM Report 679102029. National Institute of Public Health and Environment, Bilthoven, The Netherlands
9. **Dixon, W.J.** and **F.J. Massey.** 1969. *Introduction to Statistical Analysis*, 3rd ed. McGraw-Hill, Kogakusha, Tokyo, Japan.
10. **Aldenberg, T.** 1993. E_TX 1.3a, A program to calculate confidence limits for hazardous concentrations based on small samples of toxicity data. RIVM Report 719102015. National Institute of Public Health and Environment, Bilthoven, The Netherlands
11. **D'Agostino, R.B.** and **M.A. Stephens.** 1986. *Goodness-of-Fit Techniques*, Vol. 68—Statistics: Textbooks and Monographs. Marcel Dekker, New York, NY, USA.

APPENDIX

List of chemicals, with data for more than three species, used for the calculation of the pooled standard deviation[a]

Compound	Number of species	Chemical family	LD50 range (mg/kg body wt.)	HD5(50) (mg/kg body wt.)
Birds				
3-Chloro-*p*-toluidine	10	m	1–422	0.41
4-Aminopyridine	34	m	1–13.3	1.76
Aldicarb	11	c	0.75–152	0.75
Aldrin	11	m	5–520	2.75
Aminocarb	4	c	22.5–212	9.12
Azinphos-methyl	6	o	8–136	7.71
Camphechlor	12	m	19.9–581	13.3
Carbofuran	19	c	0.24–38.9	0.29
Carbophenothion	8	o	5.62–316	3.42
Chlordane	4	m	14.1–1,200	2.31
Chlorfenvinphos	13	o	3.16–240	5.16
Chlormequat	4	m	261–920	134
Chlorpyrifos	17	o	5–157	7.57
Demeton	12	o	1.33–39	1.43
Diazinon	12	o	1.8–316	0.98
Dichlofenthion	8	o	14–2,370	13.6
Dichlorvos	9	o	11–42.2	8.23
Dicrotophos	14	o	1–10	1.11
Dieldrin	13	m	8.78–381	6.37
Dimethoate	7	o	6.6–152	6.03
Dinoseb	5	m	7.1–27	4.54
Disulfoton	6	o	2.37–31.6	1.31
Endosulfan	4	m	6.47–320	8.31
Endrin	11	m	0.316–22.3	0.84
EPN	14	o	2.37–274	1.09
Ethoprophos	9	o	4.2–61	2.57
Fenitrothion	11	o	11–2,550	5.45
Fensulfothion	13	o	0.24–40	0.18
Fenthion	22	o	1–40	1.37
Fonofos	8	o	10–128	6.99
Gophacide	7	m	2.5–322	1.20
Heptachlor	5	m	62.4–2,000	24.2
Isofenphos	4	o	3–33	3.66
Malathion	5	o	167–1,485	109
Methiocarb	31	c	1.33–1,000	1.66
Methomyl	9	c	10–42.2	9.01
Methylparathion	7	o	3.08–60.5	2.77
Mevinphos	11	o	1.13–23.7	0.70
Mexacarbate	9	c	2.37–50.4	1.81
Monocrotophos	19	o	0.19–24.3	0.41
Paraquat	4	m	199–4,048	141
Parathion	19	o	0.13–160	0.47
Phorate	8	o	0.63–21	0.59
Phosphamidon	12	o	1.5–11.8	1.73
Pirimicarb	6	c	8.2–54	7.93
Propoxur	23	c	3.55–750	2.46
Sodium monofluoroacetate	10	m	3–12.8	2.41
Starlicide	30	m	1–562	0.55
Strychnine	8	m	2–112	1.90
Temephos	11	o	18.9–1,000	14.8
Thalliumsulphate	4	m	23.7–120	15.4
Trichlorfon	9	o	22.4–123	22.0
Trichloronat	8	o	1.6–1,000	1.13
Trimethacarb	7	c	10–168	13.0
Zincphosphide	6	m	13.5–237	8.08
Mammals				
1,1,1-Trichloroethane	4	m	5,660–12,300	4,767
1,2-Dibromoethane	4	m	55–420	26.7
2,4-D	5	m	100–1,200	185
Acephate	4	o	215–2,025	118
Aldicarb	4	c	0.3–1.3	0.39
Aldrin	5	m	10–95	25.9
Atrazine	4	m	250–4,080	139
Barban	4	c	240–1,500	153
Bendiocarb	5	c	28–156	17.1
Bentazone	4	m	400–1,130	416
Brodifacoum	6	m	0.25–50	0.018
Bromophos-ethyl	4	o	28–550	13.8
Camphechlor	9	m	15–375	24.3
Carbaryl	8	c	108–2,000	141
Carbofuran	6	c	1.7–34.5	2.60
Carbophenothrion	4	o	7–1,250	3.65
Chloralose	4	m	100–1,000	61.2
Chlordane	5	m	20–1,000	61.4
Chlordimeform	4	m	100–625	70.1
Chloridazon	4	m	500–3,830	600
Chlorpyrifos	4	o	82–2,000	43.6
Chlorpyrifos-methyl	4	o	1,090–3,733	1,765
Dazomet	4	m	120–640	108
Diazinon	5	o	70–480	57.7
Dichlobenil	4	m	200–4,500	103
Dichlorophen	4	m	1,000–2,600	693
Dieldrin	14	m	24–210	27.5
Dimethoate	9	o	28–680	51.3
Dinocap	4	m	53–2,000	10.9
Diquat	6	m	30–440	30.8
DNOC	8	m	16–4,200	14.8
Edifenphos	4	o	63–670	135
Endrin	4	m	3–43.4	3.46
EPN	4	o	7.7–91	15.1
EPTC	4	c	112–3,160	59.5
Ethiofencarb	5	c	72–500	70.6
Fenamidophos	6	o	2.3–100	3.2
Fenthion	5	o	40–1,000	30.7
Fentin-acetate	4	m	10–491	8.3
Ferroussulfate	6	m	300–5,000	265
Formothion	4	o	83.3–540	101
Heptachlor	5	m	55–220	62.9
Hexachlorobenzene	4	m	1,700–4,000	1,354
Isobenzan	7	m	2.3–20	1.37
Lindane	4	m	25–562	26.4
Malathion	4	o	80–12,500	127
Mecarbam	5	c	20–106	16.6
Metaldehyde	5	m	100–1,250	168
Methamidophos	6	o	10–32.2	12.4
Methidathion	6	o	17–200	10.8
Methiocarb	4	c	23–100	11.1
Methomyl	4	c	10–48	6.35
Omethoate	5	o	23–100	24.3
Oxydemeton-methyl	4	o	30–120	20.0
Paraquat	10	m	22–260	21.8
Parathion	8	o	1.75–56	2.68
PCP	6	m	27–300	85.0
Phenthoate	6	o	72–4,728	74.6
Pirimiphosmethyl	6	o	575–2,300	97
Pyriminil	12	m	4.75–4,000	4.56
Silvex	4	m	600–1,410	460
Sodium fluoroacetate	5	m	0.06–2.5	0.046
TCA	5	m	3,200–6,000	3,143
Thiodicarb	5	c	39–556	50.5
Thiometon	4	o	62–261	36.8
Thiophanate-methyl	5	c	2,250–7,500	1,854
Thiram	5	c	210–4,000	69.2
Trichlorfon	6	o	100–1,370	78.3
Trichloronat	5	o	10–100	12.9

[a] Of each compound the number of test species, the range of the LD50 values is presented and the chemical family to which the chemical belongs. In addition the median estimate of the hazardous dose for 5% of the species is given (HD5[50]). c = Carbamates; o = organophosphorus; m = miscellaneous.